(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 335 739 A2

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 89303217.7

(22) Date of filing: 31.03.89

(51) Int. Cl.[4]: G 10 L 7/00
G 06 K 9/64

(30) Priority: 31.03.88 JP 78827/88

(43) Date of publication of application:
04.10.89 Bulletin 89/40

(84) Designated Contracting States: DE FR GB NL

(71) Applicant: KABUSHIKI KAISHA TOSHIBA
72, Horikawa-cho Saiwai-ku
Kawasaki-shi Kanagawa-ken 210 (JP)

(72) Inventor: Segawa, Hideo c/o Kabushiki Kaisha Toshiba
Patent Div. 1-1 Shibaura 1-chome Minato-ku
Tokyo 105 (JP)

(74) Representative: Freed, Arthur Woolf et al
MARKS & CLERK 57-60 Lincoln's Inn Fields
London WC2A 3LS (GB)

(54) Pattern recognition system.

(57) A pattern recognition system includes a feature extracting section (4) for extracting a feature of an input pattern, a memory section (8) for storing a reference pattern for each category, a similarity calculation section (7) for calculating a similarity between the feature obtained by the feature extracting section and the reference pattern stored in the memory section (8), and a posterior probability transformation section (9) for transforming the similarity calculated by the similarity calculation section (7) into a posterior probability. The posterior probability transformation section (9) calculates the posterior probability calculated by using a parameter set required for calculating the posterior probability and calculated in recognition processing of each category in advance on the basis of the similarity calculated by the similarity calculation section (7) and a category thereof.

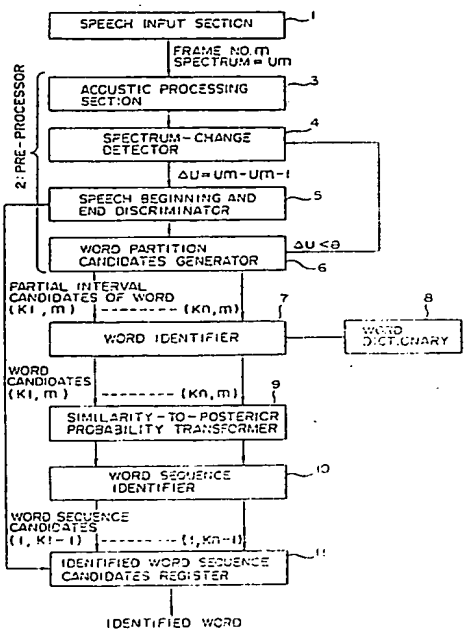

FIG. 4

EP 0 335 739 A2

BEST AVAILABLE COPY

Description

Pattern recognition system

The present invention relates to a pattern recognition system capable of improving recognition accuracy by combining posterior probabilities obtained from similarity values (or differences between reference patterns and input patterns) of input acoustic units or input characters in pattern recognition such as speech recognition or character string recognition and, more particularly, to a pattern recognition system in which an apriori probability based on contents of a lexicon is reflected in a posterior probability.

Known conventional pattern recognition systems recognize continuously input utterances or characters in units of word or character sequences. As one of such pattern recognition systems, a connected digit speech recognition algorithm using a method called a multiple similarity (MS) method will be described below.

Continuously uttered input utterances in a system are divided into frames of predetermined times. For example, an input utterance interval [1, m] having lst to m-th frames as shown in Fig. 1 will be described. In preprocessing of speech recognition, a spectral change is extracted each time one frame of an utterance is input, and word boundary candidates are obtained in accordance with the magnitude of the spectral changes. That is, a large spectral change can be considered a condition of word boundaries. In this case, the term "word" means a unit of an utterance to be recognized. The referred speech recognition system is composed of a hierarchy of lower to higher recognition levels, e.g., a phoneme level, a syllable level, a word level and a sentence level. The "words" as units of utterances to be recognized correspond to a phoneme, a syllable, a word, and a sentence at the corresponding levels. Word recognition processing is executed whenever the word boundary candidate is obtained.

In the word sequence recognition processing, the interval [1, m] is divided into two partial intervals, i.e., intervals [1, ki] and [ki, m]. ki indicates the frame number of the i-th word boundary candidate. The interval [1, ki] is an utterance interval corresponding to a word sequence $\overline{wi}$, and the interval [ki, m] is a word utterance interval corresponding to a single word wi. A word sequence Wi is represented by:

$$Wi = \overline{wi} + wi \quad ...(1)$$

and corresponds to a recognition word sequence candidate of the utterance interval [1, m] divided by the i-th frame. The recognition word sequence candidates Wi are obtained for all the word boundary candidates ki ($i = 1, 2, ..., \ell$). Of these candidates thus obtained, a word sequence W having a maximum similarity value (value representing a similarity of this pattern with respect to a reference pattern) is adopted as a recognition word sequence of the utterance interval [1, m]. Note that ℓ represents the number of recognition word sequence candidates corresponding to partial intervals to be stored upon word sequence recognition and is a parameter set in the system. By sequentially increasing m by this algorithm, recognition word sequences corresponding to all the utterance intervals can be obtained.

In the above continuous speech recognition method, the number of input words is unknown. Therefore, in order to correctly recognize an input utterance pattern L as a word sequence W, whether each detected interval correctly corresponds to an uttered word must be considered. Even if this is considered, it is difficult to obtain a high recognition rate in the word sequence recognition as long as the similarity values are merely combined. This is because the similarity is not a probabilistic measure.

Therefore, some conventional systems transform an obtained similarity value into a posterior probability and use this posterior probability as a similarity measure for achieving higher accuracy than that of the similarity.

Assume that speech recognition is to be performed for an input word sequence

$W = w1\ w2 \ldots wn\ wi \in C$

including n words belonging to word set $C = \{c1, c2, ..., cN\}$ so as to satisfy the following two conditions:

(1) A word boundary is correctly recognized.

(2) The word category of each utterance interval is correctly recognized.

In this case, as shown in Fig. 2, assume that each word wi corresponds to a pattern ℓi in each partial utterance interval to satisfy the following relation:

$L = \ell 1\ \ell 2 \ldots \ell n$

In this case, if the word sequence W has no grammatical structure, wi and wj can be considered independent events ($i \neq j$). Hence the probability that each utterance interval is correctly recognized to be a corresponding word is represented by the following equation:

$$P(W \mid L) = \prod_{i=1}^{n} P(wi \mid \ell i) \quad ...(2)$$

In this equation, P(W | L) is called likelihood. Upon calculation of the P(W | L), in order to prevent repetition of multiplication, logarithms of both sides of equation (2) are often taken to obtain logarithmic likelihood as follows:

$$\log P(W \mid L) = \sum_{i=1}^{n} \log P(wi \mid \ell i) \quad ...(3)$$

In this equation, P(wi | ℓi) is a conditional probability that an interval ℓi corresponds to wi and is a posterior probability to be obtained.

Therefore, by transforming an obtained similarity value into a posterior probability by a table, a high recognition rate can be obtained.

Since it is practically difficult to obtain the posterior probability P(wi | ℓi), however, a similarity value is normally used instead of a probability value, while properly biasing the similarity value to make it approximate to a probability value. For example, Ukita et al. performed approximation by an exponential function as shown in Fig. 3 ("A Speaker Independent Recognition Algorithm for Connected Word Boundary Hypothesizer, Proc. ICASSP, Tokyo, April, 1986):

$$\log P = \begin{cases} S - Smax & (S < Smax) \\ 0 & (otherwise) \end{cases} \quad \ldots(5)$$

A logarithm of the equation (4) is calculated and the relation $A \cdot B^{Smax} = 1.0$ is utilized to obtain the following equation:

$$P = \begin{cases} A \cdot B^{S} & (S < Smax) \\ 1 & (otherwise) \end{cases} \quad \ldots(4)$$

By subtracting a fixed bias Smax from similarity S, a similarity value is transformed into a probability value. When this measure is used in connected digit speech recognition, the bias Smax is set to be 0.96.

A posterior probability curve, however, is not generally a fixed curve but a variable one depending on a size of a lexicon or the contents of the lexicon (e.g., the number of similar words is large). Therefore, the conventional method of transforming a similarity value into a posterior probability on the basis of only one fixed curve as described against many applications cannot perform recognition with high accuracy.

As described above, in the conventional pattern recognition system for estimating similarity by transforming the similarity into a posterior probability, a transformation curve for obtaining the posterior probability is approximated to a fixed curve because it is difficult to obtain a curve corresponding to the contents of a lexicon or the number of words. Therefore, recognition cannot be performed with high accuracy.

It is an object of the present invention to provide a pattern recognition system capable of performing recognition with high accuracy by performing similarity-posterior transformation on the basis of a parameter easily obtained by learning the training data belonging to the lexicon.

The pattern recognition system according to the present invention performs posterior probability transformation processing for transforming a similarity value calculated from the feature vectors of an input pattern and a reference pattern for each category into a posterior probability calculated on the basis of the recognized category, the calculated similarity and a transformation parameter acquired from learning in advance.

The transformation parameter is a parameter set including parameters for defining a distribution of similarities of correctly recognized input patterns in recognition processing acquired from the similarity value training data of each category, parameters for defining a distribution of similarities of erroneously recognized input patterns in the recognition processing, and a weighting coefficient ω required for calculating the posterior probability from the distributions of the two parameters. In transformation calculation, the posterior probability is calculated on the basis of the similarity value calculated and the above transformation parameter set.

That is, in the pattern recognition process, predetermined calculation is performed by using transformation parameters corresponding to the recognition result, thereby transforming a similarity value into a desired posterior probability. In addition, the transformation requires complicated calculations. Therefore, by setting the calculation results into a table in advance, a processing speed can be increased.

Therefore, according to the pattern recognition system of the present invention, a correct posterior probability transformation parameter can be obtained by a small number of samples, and the accuracy of recognition processing can be greatly improved by using the parameters.

This invention can be more fully understood from the following detailed description when taken in conjunction with the accompanying drawings, in which:

Fig. 1 is a view for explaining an input utterance interval and a word boundary candidate;

Fig. 2 is a view showing a correspondence between an utterance pattern and a word sequence;

Fig. 3 is a graph showing an approximate transformation function used in a conventional system;

Fig. 4 is a block diagram of a continuous speech digit recognition system according an embodiment of the present invention;

Fig. 5 is a block diagram showing an arrangement of a similarity-posterior probability transformation section of the system shown in Fig. 4;

Fig. 6 is a flow chart showing parameter learning steps of the similarity-posterior probability transformation section of the system shown in Fig. 4; and

Fig. 7 is a graph showing posterior probability curves obtained in the parameter learning steps of the

BEST AVAILABLE COPY

similarity-posterior probability transformation section of the system shown in Fig. 4.

A principle of a pattern recognition system according to an embodiment of the present invention will be described below.

In the system of the present invention, posterior probability transformation processing for transforming a similarity value calculated from a feature vector of an input pattern and a reference pattern for each category into a posterior probability is typically performed by a transformation parameter memory section and a transformation calculation section to be described below or by a table having functions of the two sections.

The transformation parameter memory section stores, in units of categories, a parameter set including parameters (α, β) for defining a distribution of similarity values correctly recognized in recognition processing derived from the similarity value training data parameters $(\bar{\alpha}, \bar{\beta})$ for defining a distribution of similarities erroneously recognized in the recognition processing, and a weighting coefficient ω required for calculating a posterior probability from the distributions of the two parameters.

The transformation calculation section calculates a posterior probability on the basis of a similarity value and the parameter set stored in the above transformation parameter memory section.

Assuming that a partial utterance pattern ℓi is classified into a word recognition result as its category and a similarity value (especially, a multiple similarity), a posterior probability $P(wi \mid \ell i)$ is rewritten as follows:

$$P(wi \mid \ell i) \rightarrow P(wi \mid Ti \wedge si) \quad (6)$$

(where Ti is an event in which a recognized category of ℓi in a multiple similarity method is wi, and si is a multiple similarity value of ℓi concerning a word wi)

Relation (6) can be transformed as follows by using the Bayes' theorem:

$$\begin{aligned} P(wi \mid Ti \wedge si) &= \frac{P(si \mid Ti \wedge wi) \cdot P(Ti \wedge wi)}{P(Ti \wedge wi)} \\ &= P(si \mid Ti \wedge wi) \cdot P(Ti \wedge wi) / \\ &\quad \{P(si \mid Ti \wedge wi) P(Ti \wedge wi) \\ &\quad + P(si \mid Ti \wedge \overline{wi}) P(Ti \wedge \overline{wi})\} \\ &= P(si \mid Ti \wedge wi) / \{P(si \mid Ti \wedge wi) \\ &\quad + P(si \mid Ti \wedge wi) \frac{P(Ti \wedge \overline{wi})}{P(Ti \wedge wi)}\} \end{aligned} \quad \ldots (7)$$

where $\overline{wi}$ is an event in which a pattern ℓi does not belong to the category wi.

Statistics in the equation (7) will be described below.

$P(si \mid Ti \wedge wi)$ will be described first.

$P(si \mid Ti \wedge wi)$ is a probability that an event in which a recognized category obtained in the multiple similarity method is wi and the category of input data is wi occurs. This curve can be approximated by the following equation:

$$P(si \mid Ti \wedge wi) = \frac{(1 - si)^{\alpha - 1}}{\beta^{\alpha} \cdot \Gamma(\alpha)} \exp\left\{-\frac{(1 - si)}{\beta}\right\} \quad \ldots (8)$$

where α and β are parameters obtained from training data: α represents the number of components not involved in the reference pattern in the multiple similarity method; and β, its mean variance. In this parameter estimation method, as described in "Distribution of Similarity Values in Multiple Similarity Method" by Hideo Segawa et al. (Shingaku Giho PRU87-18, June 1987), and an effective amount of training data for parameter estimation is only several tens of samples.

$P(si \mid Ti \wedge \overline{wi})$ will be described below.

$P(si \mid Ti \wedge \overline{wi})$ is a probability that an event in which a recognized category in the multiple similarity method is wi while the category of input data is not wi occurs. In continuous speech recognition, especially $\overline{wi}$ is problematic. Therefore, not only a combination of categories which easily causes $\overline{wi}$ to be erroneously recognized as wi, but also word contexts which are patterns not corresponding to a particular category involved in the lexicon and easily causing erroneous recognition such as:

(1) Part of a certain word

(Ex) "6 [roku]"

→ "6-9 [roku-kyuu]"

(2) Transient part between words

(Ex) "3-1 [san-ichi]"

→ "3-2-1 [san-ni-ichi]"

(3) Combination of two word patterns

(Ex) "2-2 [ni-ni]"

→ "2 [ni]" must be examined, and their similarity distributions must be estimated. (Within the brackets are phonetic symbols indicating how the numerals are pronounced in the Japanese language.) The similarity distribution can be approximated by the equation (8). Parameters in this similarity distribution are $(\overline{\alpha i}, \overline{\beta i})$ so as to be distinguished from the parameters $(\alpha i, \beta i)$ in the equation (8). The parameters $(\overline{\alpha i}, \overline{\beta i})$ can be easily calculated similarly to the parameters $(\overline{\alpha i}, \overline{\beta i})$.

$P(Ti \wedge wi)/P(Ti \wedge \overline{wi})$ will be taken into consideration. This statistic corresponds to an apriori probability in the Bayes' probability and to an occurrence frequency ratio of a category. $P(Ti \wedge wi)$ represents a probability that an event in which a recognition result obtained by a subspace method is wi and an input pattern is wi occurs. This statistic is calculated in a learning procedure as follows:

$$P(Ti \wedge wi)/P(Ti \wedge \overline{wi})$$

$$= \frac{(\text{frequency that events Ti and wi occur})}{(\text{frequency that events Ti and wi occur})}$$

$$= \omega$$

The obtained ω is a weighting coefficient.

As described above, each parameter set is a statistic which can be easily calculated by learning.

In pattern recognition, a set of necessary parameters α, β, $\overline{\alpha}$, $\overline{\beta}$ and ω are read out from the transformation parameter memory section in accordance with the obtained similarity si to perform a calculation based on the equation (7) in the transformation calculation section, thereby transforming the similarity value into a desired posterior probability. The transformation calculation section must perform complicated calculations. Therefore, by setting the results of transformation calculation into a table, a processing speed can be further increased.

As a result, a posterior probability transforming means can be constituted by a small data amount with high accuracy, thereby improving recognition accuracy.

A word sequence recognition system as the pattern recognition system according to the embodiment of the present invention based on the above principle will be described below.

Fig. 4 shows an arrangement of the word sequence recognition system for connected digit speech recognition.

Referring to Fig. 4, an utterance input section 1 transforms a continuous utterance into a predetermined electrical signal and supplies the signal to a preprocessor 2. The preprocessor 2 comprises an acoustic process section 3, a spectral change extraction section 4, an utterance start/end point determination section 5, and a word boundary candidate generation section 6. The acoustic process section 3 performs spectral analysis for the input utterance data in units of frames by using a filter bank of, e.g., 8 to 30 channels, thereby extracting a feature pattern. The spectral change extraction section 4 extracts a difference ΔU between spectrum data Um of each frame. The utterance start/end point determination section 5 detects the start and end points of the utterance on the basis of the magnitude of the extracted spectral change. When the spectral change ΔU is larger than a predetermined threshold value θ, the word boundary candidate generation section 6 outputs the corresponding frame as a word boundary candidate ki.

The feature patterns corresponding to n word interval candidates [ki, m] obtained by the boundary candidates ki (i = 1 to n) are supplied to a word recognition section 7 and subjected to word recognition using a word dictionary 8 therein. A word recognition candidate of each word interval candidate is transformed into a posterior probability by a similarity-posterior probability transformation section 9 and supplied to a word sequence recognition section 10. The word sequence recognition section 10 combines a word sequence candidate for each word sequence interval [1, ki] (i = 1 to k) registered in a recognition word sequence candidate registration section 11 with the similarity transformed into the posterior probability to perform word sequence recognition. Word sequence recognition candidates thus obtained are stored in the recognition word sequence candidate registration section 11. When the utterance start/end point determination section 5 detects the end point of the utterance, one of the registered word sequence candidates having a highest similarity is output as a recognized word.

Fig. 5 shows an arrangement of the similarity-posterior probability transformation section 9. The section 9 comprises a transformation calculation section 21 and a transformation parameter memory section 22. The transformation parameter memory section 22 is a table which stores parameters such as:

α, β similarity distribution of correct patterns
$\overline{\alpha}$, $\overline{\beta}$ similarity distribution of incorrect patterns
ω apriori probability ratio of correct pattern to incorrect pattern

These parameter sets can be calculated by learning.

Fig. 6 shows an algorithm of this learning.

That is, this learning processing includes first and second learning steps 31 and 32. In the first learning step 31, uttered word sequence data is divided or classified in accordance with word boundary data and a word category given as instructive data to form a reference pattern (template) of a word utterance based on the

BEST AVAILABLE COPY

multiple similarity method. In the second learning step 32, a word sequence is uttered again in accordance with the word boundary data and the word category given as the instructive data to generate a word utterance interval candidate, and a word similarity calculation with respect to the reference pattern (template) formed in the above first learning step is performed on the basis of the generated word interval candidate data, thereby obtaining word similarity data and a word recognition result. On the basis of the result and the given instructive data, correct and incorrect data similarity distributions and a category appearance frequency are calculated to obtain a posterior probability curve concerning a similarity value.

The posterior probability curve obtained as a result of the above learning is shown in Fig. 7.

When the learning is performed for all categories, parameters (αi, βi, $\overline{\alpha i}$, $\overline{\beta i}$, ω) for all the categories can be obtained. These parameters are stored in the transformation parameter memory section 22.

The transformation calculation section 21 transforms similarities into the following equations:

$$P(si \mid Ti \wedge wi) = \frac{(1 - si)^{\alpha-1}}{\beta i^{\alpha i} \cdot \Gamma(\alpha i)} \exp\left\{ - \frac{(1 - si)}{\beta i} \right\} \quad \ldots(9)$$

$$P(si \mid Ti \wedge \overline{wi}) = \frac{(1 - si)^{\overline{\alpha}-1}}{\beta i^{\overline{\alpha i}} \cdot \Gamma(\overline{\alpha i})} \exp\left\{ - \frac{(1 - si)}{\overline{\beta i}} \right\} \quad \ldots(10)$$

and then calculates the posterior probability by the following transformation equation:

$$P(wi \mid Ti \wedge si) = P(si \mid Ti \wedge wi)/\{P(si \mid Ti \wedge wi) + P(si \mid Ti \wedge \overline{wi}) \cdot \omega\} \quad (11)$$

As described above, according to the system of the present invention, the similarity-posterior probability transformation section can be easily formed by the simple learning processing, and the recognition processing can be performed with high accuracy by using the obtained transformation section.

Upon transformation into a posterior probability, different transformation curves are preferably used for the respective recognition categories. Since a common transformation curve is used depending on the recognition category result, however, the following equation may be used:

$$P = \sum_{i=1}^{N} P(Wi \mid Ti \wedge si)/N \quad \ldots(12)$$

In addition, since the transformation calculation section must perform complicated calculations, the transformation calculation section and the transformation parameter memory section may be combined into a table. As a result, a transformation speed can be increased.

The present invention can be applied to not only speech recognition but also another pattern recognition such as character recognition.

Claims

1. A pattern recognition system comprising:
feature extracting means (4) for extracting a feature of an input pattern;
memory means (8) for storing a reference pattern for each category;
similarity calculating means (7) for calculating a similarity between the feature vector extracted by said feature extracting means (4) and the reference pattern stored in said memory means (8); and
posterior probability transforming means (9) for transforming the similarity calculated by said similarity calculating means (7) into a posterior probability,
characterized in that said posterior probability transforming means includes means for calculating the posterior probability by using a parameter set required for calculating the posterior probability on the basis of the similarity calculated by said similarity calculating means and a category thereof, said parameter set obtained in recognition processing from training similarity value data in advance.

2. A pattern recognition system comprising:
feature extracting means (4) for extracting a feature of an input pattern;
memory means (8) for storing a reference pattern of each category;
similarity calculating means (7) for calculating a similarity values between the feature vector extracted by said feature extracting means (4) and the reference pattern stored in said memory means (8); and
posterior probability transforming means (9) for transforming the similarity value calculated by said

BEST AVAILABLE COPY

similarity calculating means (7) into a posterior probability,
characterized in that said posterior probability transforming means includes:
transformation parameter memory means (22) for storing, for each category, a parameter set including parameters for defining a distribution of similarity values correctly recognized in recognition processing derived from training similarity value in advance, parameters for defining a distribution of similarities erroneously recognized in the recognition processing, and a weighting coefficient required for calculating the posterior probability from the two parameter distributions; and
calculating means (21) for calculating the posterior probability on the basis of the similarity calculated by said similarity calculating means (7) and the parameter set stored in said transformation parameter memory means.

3. A system according to claim 2, characterized in that said transformation parameter memory means (22) includes means for storing, for each category, a parameter set including parameters α and β for defining the distribution of similarities correctly recognized in the recognition processing derived from the training similarity value data in advance, parameters $\overline{\alpha}$ and $\overline{\beta}$ for defining the distribution of similarities erroneously recognized in the recognition processing, and a weighting coefficient ω required for calculating the posterior probability from the two parameter distributions.

4. A system according to claim 3, characterized in that said calculating means (21) includes means for transforming a similarity into the following equations on the basis of the similarity calculated by said similarity calculating means (7) and the parameter set including α, β, $\overline{\alpha}$, $\overline{\beta}$ and ω stored in said transformation parameter memory means (22):

$$P(si \mid Ti \wedge wi) = \frac{(1 - si)^{\alpha - 1}}{\beta i^{\alpha i} \cdot \Gamma(\alpha i)} \exp\left\{ - \frac{(1 - si)}{\beta i} \right\}$$

$$P(si \mid Ti \wedge \overline{wi}) = \frac{(1 - si)^{\overline{\alpha} - 1}}{\overline{\beta i}^{\alpha i} \cdot \Gamma(\overline{\alpha i})} \exp\left\{ - \frac{(1 - si)}{\overline{\beta i}} \right\}$$

and calculating the posterior probability by the following transformation equation:

$$P(wi \mid Ti \wedge si) = P(si \mid Ti \wedge wi) / \{P(si \mid Ti \wedge wi) + P(si \mid Ti \wedge \overline{wi}) \cdot \omega\}.$$

5. A pattern recognition system comprising:
feature extracting means (4) for extracting a feature of an input pattern;
memory means (8) for storing a reference pattern for each category;
similarity calculating means (7) for calculating a similarity between the feature extracted by said feature extracting means (4) and the reference pattern stored in said memory means (8); and
posterior probability transforming means (9) for transforming the similarity calculated by said similarity calculating means (7) into a posterior probability,
characterized in that said posterior probability transforming means includes table memory means for storing a transformation table for calculating the posterior probability calculated by using a parameter set including a parameter for defining a distribution of similarity values correctly recognized in recognition processing derived from the training similarity value data in advance with respect to the reference pattern of each category on the basis of the similarity calculated by said similarity calculating means (7) and a category thereof, a parameter for defining a distribution of similarities erroneously recognized in the recognition processing, and a weighting coefficient required for calculating the posterior probability from the two parameter distributions, and the similarity calculated by said similarity calculating means (7).

6. A system according to claim 5, characterized in that said table memory means (9) includes means for storing, for each category, a posterior probability calculated by using the parameter set including parameters α and β for defining the distribution of similarity values correctly recognized in the recognition processing derived from training similarity value data of each category, parameters $\overline{\alpha}$ and $\overline{\beta}$ for defining the distribution of similarities erroneously recognized in the recognition processing, and a weighting coefficient ω required for calculating the posterior probability from the two parameter distributions, and the similarity calculated by said similarity calculating means (7).

7. A system according to claim 6, characterized in that said table memory means (9) includes means for transforming a similarity into the following equations on the basis of the similarity calculated by said similarity calculating means (7) and the parameter set including α, β, $\overline{\alpha}$, $\overline{\beta}$ and ω stored in said transformation parameter memory means:

BEST AVAILABLE COPY

$$P(si \mid Ti \wedge wi) = \frac{(1 - si)^{\alpha - 1}}{\beta i^{\alpha i} \cdot \Gamma(\alpha i)} \exp\left\{ - \frac{(1 - si)}{\beta i} \right\}$$

$$P(si \mid Ti \wedge \overline{wi}) = \frac{(1 - si)^{\overline{\alpha} - 1}}{\overline{\beta i}^{\alpha i} \cdot \Gamma(\overline{\alpha i})} \exp\left\{ - \frac{(1 - si)}{\overline{\beta i}} \right\}$$

and storing the posterior probability calculated by the following transformation equation:

$P(wi \mid Ti \wedge si) = P(si \mid Ti \wedge wi)/\{P(si \mid Ti \wedge wi) + P(si \mid Ti \wedge \overline{wi}) \cdot \omega\}$.

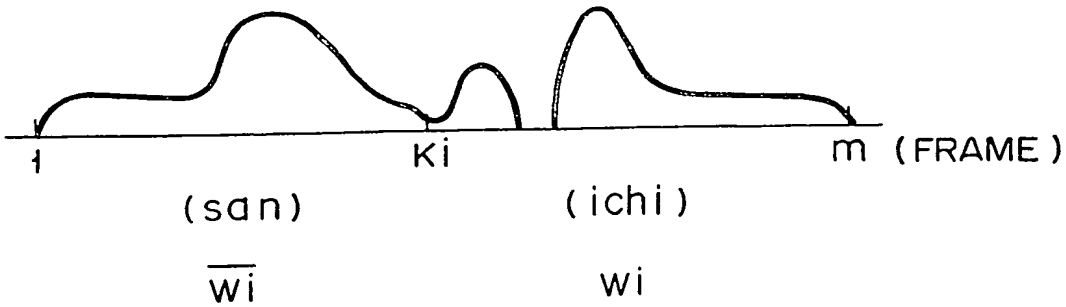

FIG. 1

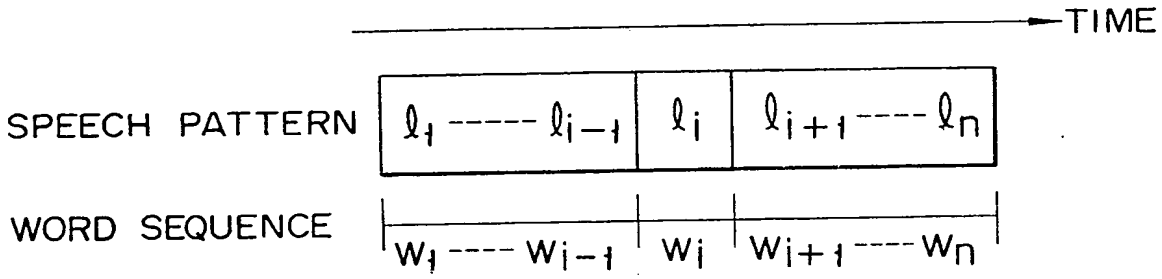

FIG. 2

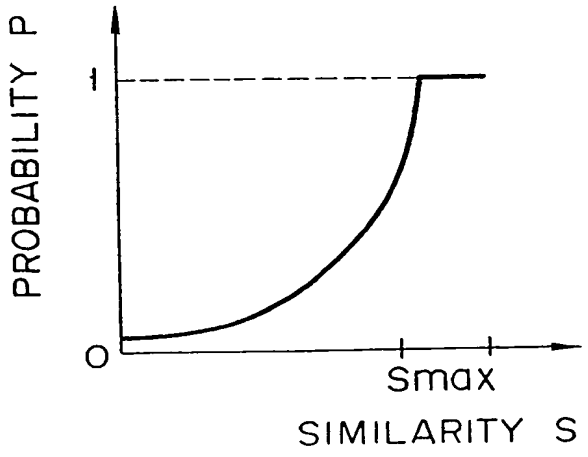

FIG. 3

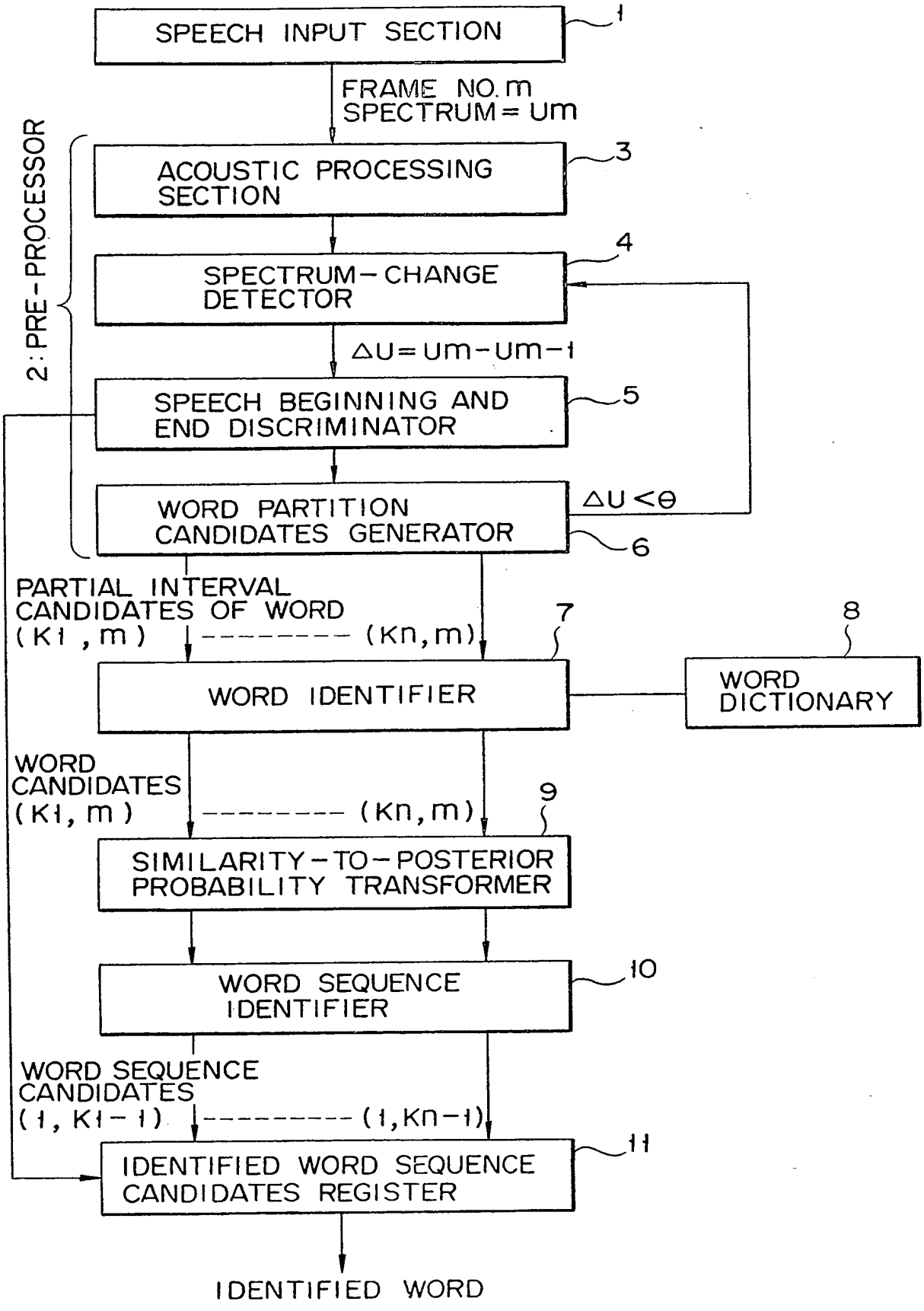
SPEECH INPUT SECTION
1
FRAME NO. m
SPECTRUM = Um
2: PRE-PROCESSOR
ACOUSTIC PROCESSING SECTION
3
SPECTRUM-CHANGE DETECTOR
4
ΔU = Um - Um-1
SPEECH BEGINNING AND END DISCRIMINATOR
5
WORD PARTITION CANDIDATES GENERATOR
ΔU < θ
6
PARTIAL INTERVAL CANDIDATES OF WORD
(K1, m) -------- (Kn, m)
7
8
WORD IDENTIFIER
WORD DICTIONARY
WORD CANDIDATES
(K1, m) -------- (Kn, m)
9
SIMILARITY-TO-POSTERIOR PROBABILITY TRANSFORMER
WORD SEQUENCE IDENTIFIER
10
WORD SEQUENCE CANDIDATES
(1, K1-1) --------- (1, Kn-1)
IDENTIFIED WORD SEQUENCE CANDIDATES REGISTER
11
IDENTIFIED WORD

FIG. 4

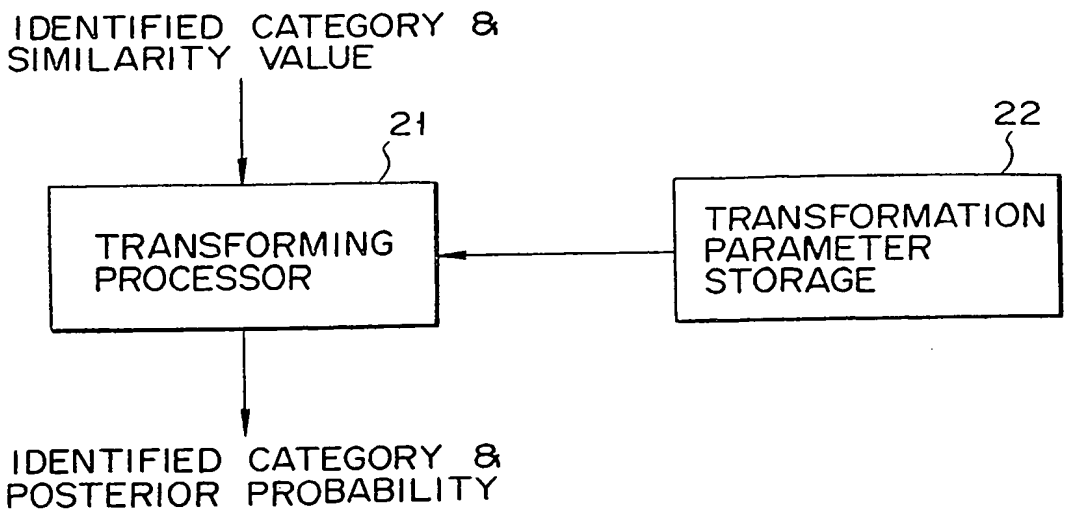
IDENTIFIED CATEGORY &
SIMILARITY VALUE
21
22
TRANSFORMING
PROCESSOR
TRANSFORMATION
PARAMETER
STORAGE
IDENTIFIED CATEGORY &
POSTERIOR PROBABILITY

FIG. 5

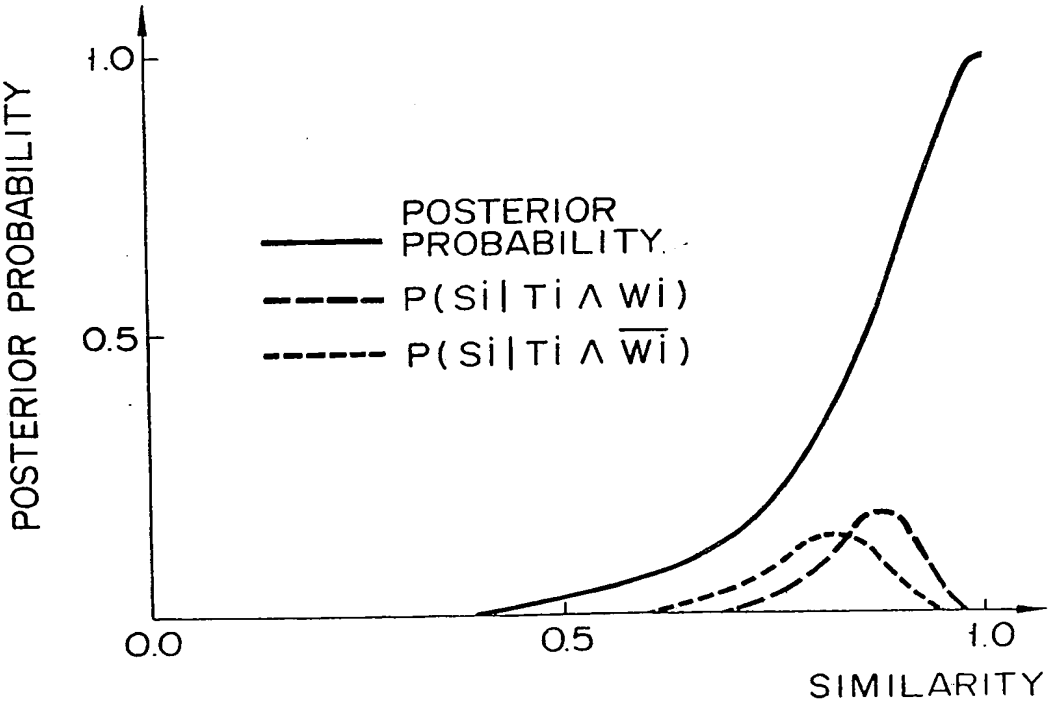
POSTERIOR PROBABILITY
1.0
0.5
0.0
0.5
1.0
SIMILARITY
POSTERIOR
PROBABILITY
P(Si | Ti ∧ Wi)
P(Si | Ti ∧ $\overline{Wi}$)

FIG. 7

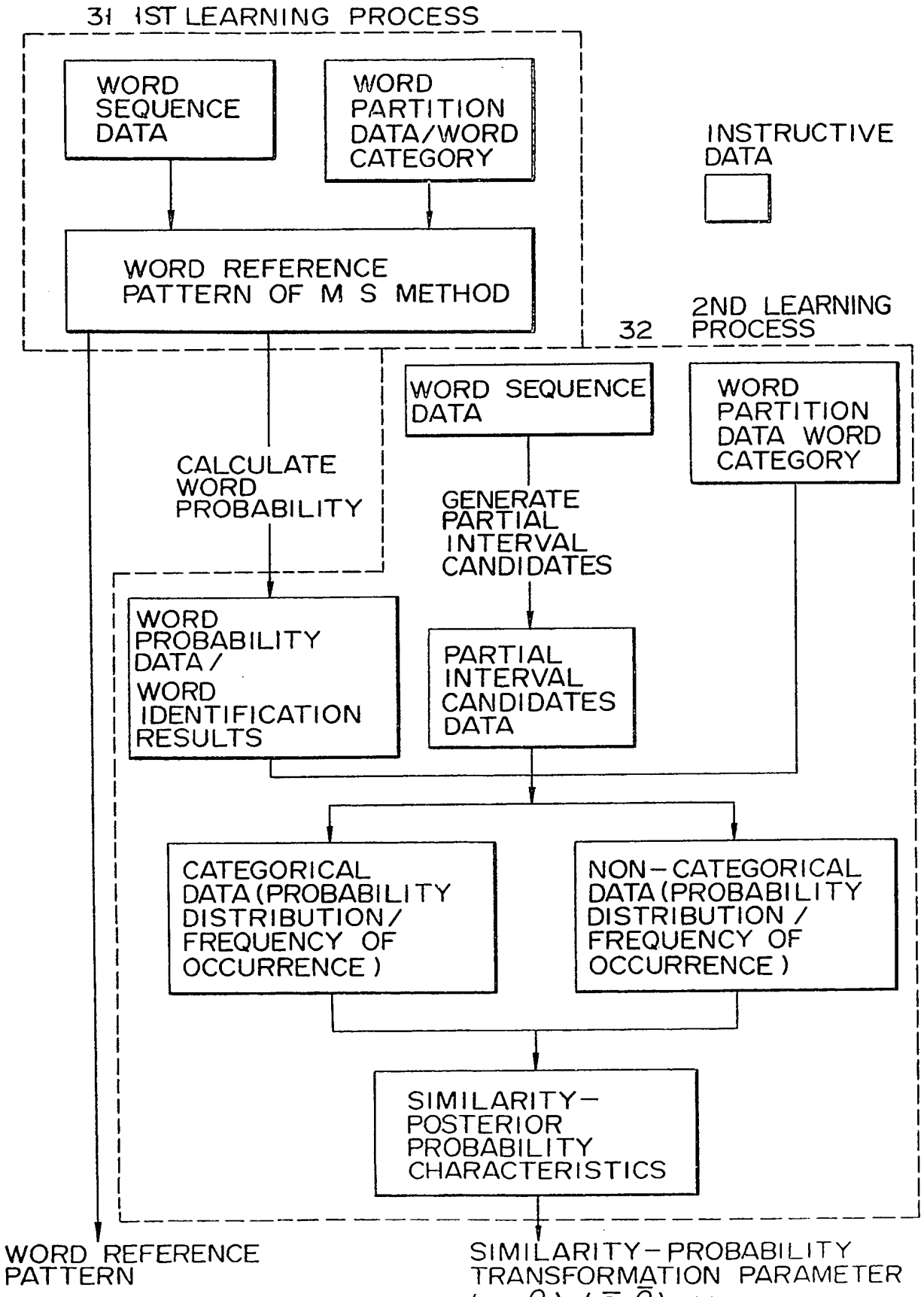
31 1ST LEARNING PROCESS
WORD SEQUENCE DATA
WORD PARTITION DATA/WORD CATEGORY
INSTRUCTIVE DATA
WORD REFERENCE PATTERN OF M S METHOD
32 2ND LEARNING PROCESS
WORD SEQUENCE DATA
WORD PARTITION DATA WORD CATEGORY
CALCULATE WORD PROBABILITY
GENERATE PARTIAL INTERVAL CANDIDATES
WORD PROBABILITY DATA / WORD IDENTIFICATION RESULTS
PARTIAL INTERVAL CANDIDATES DATA
CATEGORICAL DATA (PROBABILITY DISTRIBUTION / FREQUENCY OF OCCURRENCE)
NON-CATEGORICAL DATA (PROBABILITY DISTRIBUTION / FREQUENCY OF OCCURRENCE)
SIMILARITY-POSTERIOR PROBABILITY CHARACTERISTICS
WORD REFERENCE PATTERN
SIMILARITY-PROBABILITY TRANSFORMATION PARAMETER
$(\alpha, \beta), (\bar{\alpha}, \bar{\beta}), \omega$

FIG. 6

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 335 739 A3**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: **89303217.7**

(51) Int. Cl.5: **G10L 7/00**, G06K 9/64

(22) Date of filing: **31.03.89**

(30) Priority: **31.03.88 JP 78827/88**

(43) Date of publication of application:
04.10.89 Bulletin 89/40

(84) Designated Contracting States:
DE FR GB NL

(88) Date of deferred publication of the search report:
28.08.91 Bulletin 91/35

(71) Applicant: **KABUSHIKI KAISHA TOSHIBA**
72, Horikawa-cho Saiwai-ku
Kawasaki-shi Kanagawa-ken 210(JP)

(72) Inventor: **Segawa, Hideo c/o Kabushiki Kaisha Toshiba**
Patent Div. 1-1 Shibaura 1-chome Minato-ku
Tokyo 105(JP)

(74) Representative: **Freed, Arthur Woolf et al**
MARKS & CLERK 57-60 Lincoln's Inn Fields
London WC2A 3LS(GB)

(54) Pattern recognition system.

(57) A pattern recognition system includes a feature extracting section (4) for extracting a feature of an input pattern, a memory section (8) for storing a reference pattern for each category, a similarity calculation section (7) for calculating a similarity between the feature obtained by the feature extracting section and the reference pattern stored in the memory section (8), and a posterior probability transformation section (9) for transforming the similarity calculated by the similarity calculation section (7) into a posterior probability. The posterior probability transformation section (9) calculates the posterior probability calculated by using a parameter set required for calculating the posterior probability and calculated in recognition processing of each category in advance on the basis of the similarity calculated by the similarity calculation section (7) and a category thereof.

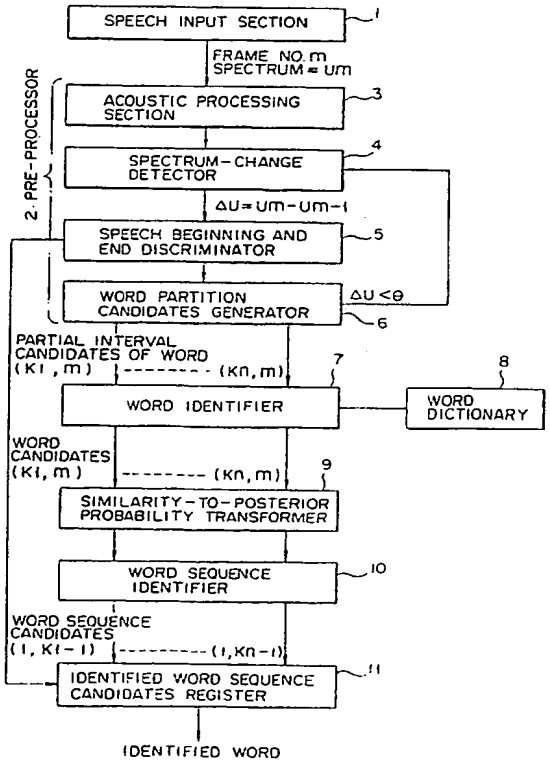

FIG. 4

EP 0 335 739 A3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 89 30 3217

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
A,D	PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP 86), TOKYO, APRIL 7 - 11 1986, IEEE, NEW YORK US; T. Ukita et al.:"A speaker independent recognition algorithm for connected word using word boundary hypothesizer", pages 1077 - 1080		G 10 L 7/00 G 06 K 9/64
P,X	PROCEEDINGS OF THE 9TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, ROME, NOVEMBER 14-17, 1988, IEEE COMPUTER SOCIETY PRESS, WASHINGTON, US; H. Segawa et al.:"A Similarity Value Transformation Method for Probabilistic Scoring", pages 1225-1229 * paragraph 3.2 *	1-7	
			TECHNICAL FIELDS SEARCHED (Int. Cl.5)
			G 06 K

The present search report has been drawn up for all claims

Place of search	Date of completion of search	Examiner
The Hague	25 June 91	SONIUS M.E.

CATEGORY OF CITED DOCUMENTS

X: particularly relevant if taken alone
Y: particularly relevant if combined with another document of the same catagory
A: technological background
O: non-written disclosure
P: intermediate document
T: theory or principle underlying the invention

E: earlier patent document, but published on, or after the filing date
D: document cited in the application
L: document cited for other reasons

&: member of the same patent family, corresponding document